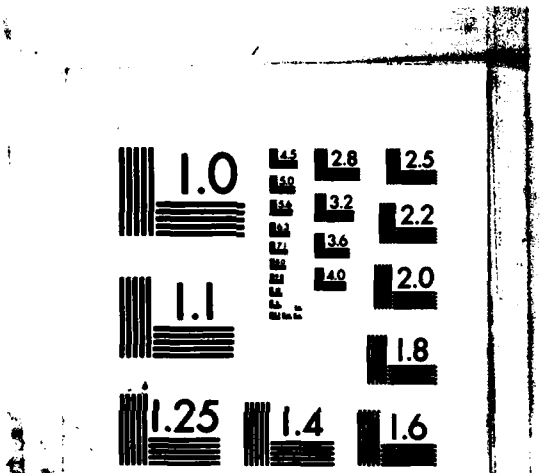
AD-A140 180 ON THE OPTIMALITY OF SOME SUBSET SELECTION PROCEDURES 1/1
(U) PURDUE UNIV LAFAYETTE IN DEPT OF STATISTICS
L Y LEU MAR 84 TR-84-7 N00014-75-C-0455

UNCLASSIFIED F/G 12/1 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

(12

AD A140180

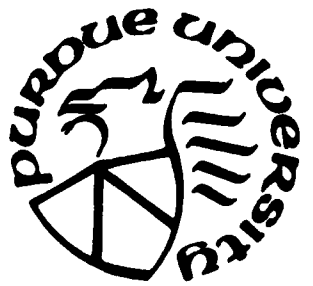
ON THE OPTIMALITY OF SOME SUBSET SELECTION PROCEDURES*

by

Lii-Yuh Leu
National Central University

Technical Report #84-7

PURDUE UNIVERSITY



DEPARTMENT OF STATISTICS

DTIC
ELECTE
APR 1 7 1984
D
A

This document has been approved for public release and sale; its distribution is unlimited

DTIC FILE COPY

84 04 16 059

ON THE OPTIMALITY OF SOME SUBSET SELECTION PROCEDURES*

by

Lii-Yuh Leu
National Central University

Technical Report #84-7

Department of Statistics
Purdue University

March 1984

This document has been approved for public release and sale; its distribution is unlimited.

*This research was supported by the Office of Naval Research Contract N00014-75-C-0455 at Purdue University. Reproduction in whole or in part is permitted for any purpose of the United States Government.

APR 17 1984

ON THE OPTIMALITY OF SOME SUBSET SELECTION PROCEDURES*

by

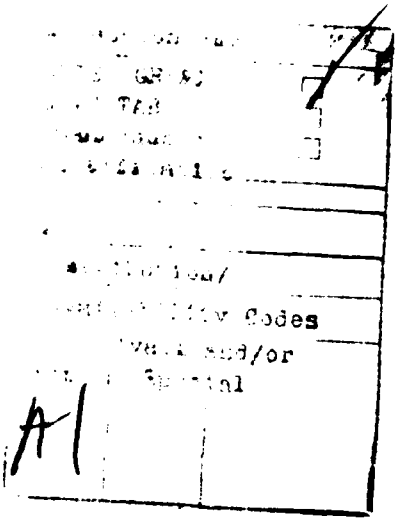
Lii-Yuh Leu
National Central University

Abstract

As measures of goodness of a selection rule, usually two quantities, the probability of a correct selection and the expected size of the selected subset, are considered. Based on these two criteria, Gupta and Huang (1980) proved a theorem to derive a selection procedure with some optimality property. However, the theorem cannot be applied to the unequal sample sizes case. In this paper, ~~we~~ its authors use a different method to generalize this theorem to the unequal sample sizes case. Also a dual problem is investigated. Also, ~~we~~ they treat a selection procedure in terms of multiple tests. Based on this approach, we derive an optimality result, is derived.

key words: Subset selection, restricted minimax, multiple tests.

*This research was supported by the Office of Naval Research Contract N00014-75-C-0455 at Purdue University. Reproduction in whole or in part is permitted for any purpose of the United States Government.



ON THE OPTIMALITY OF SOME SUBSET SELECTION PROCEDURES*

by

Lii-Yuh Leu
National Central University

1. Introduction

Let $\pi_1,\pi_2,\ldots,\pi_k$ be k independent populations which are characterized by parameters $\theta_1,\theta_2,\ldots,\theta_k$, respectively. Let X_i be the observation from population π_i, $1 \le i \le k$. Assume $\underline{X} = (X_1,\ldots,X_k)$ is a sufficient statistic for $\underline{\theta} = (\theta_1,\ldots,\theta_k)$. Suppose that we are interested in selecting a subset of the k populations containing the largest parameter (or the smallest parameter). For subset selection, let $\delta = (\delta_1,\ldots,\delta_k)$ be a selection procedure, where $\delta_i(\underline{x})$ is the probability of selecting π_i, $1 \le i \le k$, based on the observed vector $\underline{X} = \underline{x}$, and $\sum_{i=1}^{k} \delta_i(\underline{x}) \ge 1$, for all $\underline{x}$ (i.e. at least one population is selected). As measures of goodness of a selection rule, usually two quantities, the probability of a correct selection and the expected size of the selected subset, are considered. Based on these two criteria, optimal selection rules have been considered by Bahadur (1950), Eaton (1967), Lehmann (1961, 1966), Studden (1967), Nagel (1970), Spjøtvoll (1972), Alam (1973), Gupta and Huang (1977, 1980), Berger and Gupta (1980), and Bjørnstad (1981), among others.

Most of the literature on the optimality of selection rules deals with the problem when the sample sizes are all equal and is restricted to the

---

*This research was supported by the Office of Naval Research Contract N00014-75-C-0455 at Purdue University. Reproduction in whole or in part is permitted for any purpose of the United States Government.

location model. Gupta and Huang (1977, 1980) formulate the problem more generally, including location and scale cases. In this paper we will follow the notation in the papers of Gupta and Huang (1977, 1980) and generalize the theorem in Gupta and Huang (1980) to the unequal sample sizes case.

In Section 3, we use a different approach to deal with the problem of Gupta and Huang (1980) and generalize it to the unequal sample sizes case. Also a dual problem is investigated. In Section 4, we treat a selection problem as a multiple tests problem. Based on this approach, we derive an optimality result.

2. Notations and Definitions

Following the notation of Gupta and Huang (1980), let $\Omega = \{\underline{\theta} = (\theta_1,\ldots,\theta_k)\}$ be the parameter space. Let $\tau_{ij} = \tau_{ij}(\underline{\theta})$ be a measure of separation between π_i and π_j. We assume that there exists a monotonically non-increasing function h such that $\tau_{ji} = h(\tau_{ij})$. Let $\Omega_i = \{\underline{\theta} | \tau_{ij}(\underline{\theta}) \geq \Delta, \forall j \neq i, 1 \leq j \leq k\}$, $1 \leq i \leq k$, and $\Omega_0 = \Omega - \bar{\Omega}$ (an indifference zone), where $\bar{\Omega} = \bigcup_{i=1}^{k} \Omega_i$. For this problem, we assume Δ and τ_{ii} are known with $\Delta > \tau_{ii}$ for all i. Let $\tau_i = \min_{j \neq i} \tau_{ij}$, $1 \leq i \leq k$. We define $\tau^* = \max_{1 \leq \ell \leq k} \tau_\ell$. The population associated with τ^* will be called the best population. It should be pointed out that if $\underline{\theta} \in \Omega_i$, then $\tau_i > \tau_j$ for all $j \neq i$. Thus if $\underline{\theta} \in \Omega_i$, then π_i is the best population. In case of any tie(s) of the populations corresponding to τ^*, any one of the tied populations is "tagged" as the best population and selection of any subset containing this population is called a correct selection.

Let the observed sample vector be denoted by $\underline{x} = (\underline{x}_1,\ldots,\underline{x}_k)$, where $\underline{x}_i = (x_{i1},\ldots,x_{in_i})$, $1 \leq i \leq k$, $x_{i1},\ldots,x_{in_i}$ are the samples from π_i, $1 \leq i \leq k$. We define

(2.1) $S(\underline{\theta},\delta) = P_{\underline{\theta}}(CS|\delta)$

= the probability of a correct selection using rule δ,

and

(2.2) $R(\underline{\theta},\delta) = \sum_{i=1}^{k} E_{\underline{\theta}}\delta_i$

= the expected size of the selected subset using rule δ.

Let there be a suitably defined statistic Z_{ij} based on the n_i and n_j independent observations from π_i and π_j, $1 \le i,j \le k$, respectively, and suppose that for any i, the statistic $\underline{Z}_i = \{Z_{ij};\ j \neq i,\ 1 \le j \le k\}$ is invariant sufficient under a transformation group G and let $\underline{\tau}_i = \{\tau_{ij};\ j \neq i,\ 1 \le j \le k\}$ be a maximal invariant under the induced group $\bar{G}$. It is well-known that the distribution of $\underline{Z}_i$ depends only on $\underline{\tau}_i$. For example, if the observations from π_i are mutually independently distributed with unknown mean θ_i, $1 \le i \le k$ and known common variance σ^2, Z_{ij} might be $\bar{X}_i - \bar{X}_j$, where $\bar{X}_i = \sum_{\ell=1}^{n_i} X_{i\ell}/n_i$ and $\bar{X}_j = \sum_{\ell=1}^{n_j} X_{j\ell}/n_j$. For any i, let the joint density of Z_{ij}, $j \neq i$, $1 \le j \le k$, be $P_{\underline{\theta}}(\underline{z}_i) = P_{\underline{\tau}_i}(\underline{z}_i)$ with respect to some σ-finite measure μ. We note that $E_{\underline{\theta}}\delta_i(\underline{X}) = E_{\underline{\tau}_i}\delta_i(\underline{Z}_i)$.

The following definitions can be found in the paper of Alam (1973) (see also Gupta and Huang (1981)).

Definition 2.1. A measurable subset S of the sample space is called monotone non-decreasing if $\underline{x} \in S$ and $\underline{y}$ satisfies $x_i \le y_i$, $1 \le i \le k$, then $\underline{y} \in S$.

Definition 2.2. Let $P_{\underline{\theta}}(S)$ denote the probability measure of S under the conditional distribution of $\underline{X}$, given $\underline{\theta}$. The distribution is said to have stochastically increasing property (SIP) in $\underline{\theta}$ if $P_{\underline{\theta}}(S) \le P_{\underline{\theta}'}(S)$ for

every monotone non-decreasing set S and for all $\theta_i \leq \theta_i'$, $1 \leq i \leq k$.

Definition 2.3. A function $\varphi(\underline{x})$ is said to be non-decreasing if $\varphi(\underline{x}) \leq \varphi(\underline{y})$ for $x_i \leq y_i$, $1 \leq i \leq k$.

3. Optimal Subset Selection Procedures Based on Criteria $S(\underline{\theta},\delta)$ and $R(\underline{\theta},\delta)$

In the following discussion, we will assume that the density function $P_{\underline{\theta}}(\underline{z}_i) = P_{\underline{\tau}_i}(\underline{z}_i)$ defined in Section 2 has the SIP in $\underline{\tau}_i$. Let $P_{\underline{\theta}}(\underline{z}_i)$ be denoted by $P_0(\underline{z}_i)$ when $\tau_{ij} = \tau_{ii} =$ constant, $j \neq i$, $1 \leq j \leq k$, and by $P_i(\underline{z}_i)$ when $\tau_{ij} = \Delta$, $j \neq i$, $1 \leq j \leq k$. Our goal is to generalize the theorem of Gupta and Huang (1980). It should be pointed out that the proof is different from that of Gupta and Huang (1980).

First we quote two lemmas from Alam (1973) and Lehmann (1961) for completeness.

Lemma 3.1. (Alam (1973))

Let $\{P_{\underline{\theta}}\}$ be a family of distributions which has SIP in $\underline{\theta}$. Then $E_{\underline{\theta}}\varphi(\underline{X}) \leq E_{\underline{\theta}'}\varphi(\underline{X})$ for all non-decreasing integrable function $\varphi(\underline{x})$ and $\theta_i \leq \theta_i'$, $1 \leq i \leq k$. (Thus, if $P_{\underline{\theta}}$ has SIP in $\underline{\theta}$, and if $\varphi(\underline{x})$ is non-decreasing in x_j, then $E_{\underline{\theta}}\varphi(\underline{X})$ is non-decreasing in θ_j.)

Lemma 3.2. (Lehmann (1961))

Let μ and λ be two probability distributions on ω_0 and ω_1 (subsets of Ω), respectively. Let A and B be two positive constants and let δ^0 maximize the integral

$$(3.1) \qquad B\int_{\omega_0} S(\underline{\theta},\delta)d\mu(\underline{\theta}) - A\int_{\omega_1} R(\underline{\theta},\delta)d\lambda(\underline{\theta})$$

where $S(\underline{\theta},\delta)$ and $R(\underline{\theta},\delta)$ are defined by (2.1) and (2.2), respectively. Then

(i) δ^0 minimizes $\sup_{\underline{\theta}\in\omega_1} R(\underline{\theta},\delta)$ subject to $\inf_{\underline{\theta}\in\omega_0} S(\underline{\theta},\delta) \geq \gamma$ provided

$$\int_{\omega_1} R(\underline{\theta},\delta^0)d\lambda(\underline{\theta}) = \sup_{\underline{\theta}\in\omega_1} R(\underline{\theta},\delta^0) \tag{3.2}$$

and

$$\int_{\omega_0} S(\underline{\theta},\delta^0)d\mu(\underline{\theta}) = \inf_{\underline{\theta}\in\omega_0} S(\underline{\theta},\delta^0) = \gamma. \tag{3.3}$$

(ii) δ^0 maximizes $\inf_{\underline{\theta}\in\omega_0} S(\underline{\theta},\delta)$ subject to $\sup_{\underline{\theta}\in\omega_1} R(\underline{\theta},\delta) \leq \gamma'$

provided

$$\int_{\omega_1} R(\underline{\theta},\delta^0)d\lambda(\underline{\theta}) = \sup_{\underline{\theta}\in\omega_1} R(\underline{\theta},\delta^0) = \gamma' \tag{3.4}$$

and

$$\int_{\omega_0} S(\underline{\theta},\delta^0)d\mu(\underline{\theta}) = \inf_{\underline{\theta}\in\omega_0} S(\underline{\theta},\delta^0). \tag{3.5}$$

The following theorem is a generalization of Gupta and Huang (1980).

Theorem 3.3. Suppose that for any i, $p_i(\underline{z}_i)/p_0(\underline{z}_i)$ is non-decreasing in $\underline{z}_i$ and that $P_{\underline{\theta}}(\underline{z}_i)$ has the SIP. If $R(\underline{\theta},\delta^0)$ is maximized at $\tau_{ij} = \tau_{ii} =$ constant, for all i,j, where δ^0 is given by

$$\delta_i^0(\underline{z}_i) = \begin{cases} 1 & \text{if } P_i(\underline{z}_i) > c_i P_0(\underline{z}_i) \\ \lambda_i & \text{if } P_i(\underline{z}_i) = c_i P_0(\underline{z}_i) \\ 0 & \text{if } P_i(\underline{z}_i) < c_i P_0(\underline{z}_i), \end{cases} \tag{3.6}$$

$c_i (> 0)$ and λ_i are determined by $\int \delta_i^0 P_i = \gamma$, $1 \leq i \leq k$. Then $\delta^0 = (\delta_1^0,\ldots,\delta_k^0)$ minimizes $\sup_{\underline{\theta}\in\Omega} R(\underline{\theta},\delta)$ subject to $\inf_{\underline{\theta}\in\Omega} S(\underline{\theta},\delta) \geq \gamma$.

Proof. Let μ be the probability distribution which assigns probability a_i (will be determined later) to the set $\omega_i = \{\underline{\theta} | \tau_{ij} = \Delta\} \subset \Omega_i$, $1 \leq i \leq k$, and λ be the probability distribution which assigns probability one to the set $\omega_0 = \{\underline{\theta} | \tau_{ij} = \tau_{ii} = \text{constant}\}$.

Let A and B be two positive constants (will be determined later). Then

$$B\int_{\bar{\Omega}} S(\underline{\theta},\delta)d\mu(\underline{\theta}) - A\int_{\Omega} R(\underline{\theta},\delta)d\lambda(\underline{\theta})$$

$$= B\sum_{i=1}^{k} \int_{\omega_i} E_{\underline{\theta}}\delta_i d\mu(\underline{\theta}) - A\int_{\omega_0}\left(\sum_{i=1}^{k} E_{\underline{\theta}}\delta_i\right)d\lambda(\underline{\theta})$$

$$= \int \sum_{i=1}^{k} \delta_i (Ba_iP_i - AP_0)$$

which is maximized by putting $\delta_i = 1$ or 0 as $Ba_iP_i >$ or $< AP_0$. Let A, B, a_i, $1 \leq i \leq k$ be satisfied the conditions $a_i > 0$, $\sum_{i=1}^{k} a_i = 1$, and $c_i = A/a_iB$, then δ^0 defined by (3.6) maximizes $B\int_{\bar{\Omega}} S(\underline{\theta},\delta)d\mu(\underline{\theta}) - A\int_{\Omega} R(\underline{\theta},\delta)d\lambda(\underline{\theta})$. Now, by assumption $P_i(\underline{z}_i)/P_0(\underline{z}_i)$ is non-decreasing in $\underline{z}_i$, then δ_i^0 is non-decreasing in $\underline{z}_i$ and by Lemma 3.1, for any $\underline{\theta} \in \bar{\Omega}$, we have

$$S(\underline{\theta},\delta^0) = E_{\underline{\tau}_i}\delta_i^0 \geq E_{\underline{\Delta}}\delta_i^0 = \int \delta_i^0 P_i = \gamma.$$

Hence $\inf_{\underline{\theta}\in\bar{\Omega}} S(\underline{\theta},\delta^0) \geq \gamma$.

On the other hand,

$$\int_{\bar{\Omega}} S(\underline{\theta},\delta^0)d\mu(\underline{\theta}) = \sum_{i=1}^{k} a_i \int \delta_i^0 P_i = \gamma.$$

Therefore $\int_{\bar{\Omega}} S(\underline{\theta},\delta^0)d\mu(\underline{\theta}) = \inf_{\underline{\theta}\in\bar{\Omega}} S(\underline{\theta},\delta^0) = \gamma$.

Next, we have

$$\int_{\Omega} R(\underline{\theta},\delta^0)d\lambda(\underline{\theta}) = \int_{\omega_0} R(\underline{\theta},\delta^0)d\lambda(\underline{\theta}) = \sup_{\underline{\theta}\in\Omega} R(\underline{\theta},\delta^0), \text{ by assumption.}$$

The theorem follows by applying Lemma 3.2. (i).

Remark: In the theorem of Gupta and Huang (1980), $c_i = c$, $1 \le i \le k$ which is a special case of Theorem 3.3. We note that if the sample sizes are not equal, in order to satisfy the condition $\int \delta_i^0 p_i = \gamma$, $1 \le i \le k$, c_i should be different.

Furthermore, we have the following theorem, which is a dual of Theorem 3.3.

Theorem 3.4. Suppose that for any i, $p_i(\underline{z}_i)/p_0(\underline{z}_i)$ is non-decreasing in $\underline{z}_i$ and that $P_{\underline{\theta}}(\underline{z}_i)$ has the SIP. If $R(\underline{\theta},\delta^0)$ is maximized at $\tau_{ij} = \tau_{ii}$ = constant, for all i,j, where δ^0 is given by (3.6) and $c_i(> 0)$ and λ_i are determined by $\sum_{i=1}^{k} \int \delta_i^0 p_0 = \gamma'$ and $\int \delta_i^0 p_i$ is independent of i. Then $\delta^0 = (\delta_1^0,\ldots,\delta_k^0)$ maximizes $\inf_{\underline{\theta}\in\Omega} S(\underline{\theta},\delta)$ subject to $\sup_{\underline{\theta}\in\Omega} R(\underline{\theta},\delta) \le \gamma'$.

Proof. By the same argument as the proof of Theorem 3.3, we have δ^0 defined by (3.6) maximizes

$$B\int_{\bar{\Omega}} S(\underline{\theta},\delta)d\mu(\underline{\theta}) - A\int_{\Omega} R(\underline{\theta},\delta)d\lambda(\underline{\theta}).$$

Now, if $\underline{\theta} \in \omega_0$, we have

$$R(\underline{\theta},\delta^0) = \sum_{i=1}^{k} \int \delta_i^0 p_0 = \gamma', \text{ by assumption.}$$

Hence

$$\int_{\Omega} R(\underline{\theta},\delta^0)d\lambda(\underline{\theta}) = \sup_{\underline{\theta}\in\Omega} R(\underline{\theta},\delta^0) = \gamma'.$$

Furthermore,

$$\int_{\bar{\Omega}} S(\underline{\theta},\delta^0)d\mu(\underline{\theta}) = \sum_{i=1}^{k} a_i \int \delta_i^0 p_i = \int \delta_i^0 p_i, \text{ since } \int \delta_i^0 p_i \text{ is independent of i,}$$

and

$$\inf_{\underline{\theta}\in\bar{\Omega}} S(\underline{\theta},\delta^0) \ge \min_{1\le i\le k} \int \delta_i^0 p_i = \int \delta_i^0 p_i.$$

since

$$\int_{\bar{\Omega}} S(\underline{\theta},\delta^0)\,d\mu(\underline{\theta}) = \inf_{\underline{\theta}\in\bar{\Omega}} S(\underline{\theta},\delta^0).$$

The theorem follows by applying Lemma 3.2 (ii).

Remark: For equal sample sizes case, $p_i(\underline{z}_i)/p_0(\underline{z}_i)$ is independent of i. If we choose $c(>0)$ and λ_0 such that $\int \delta_i^0 p_0 = \gamma'/k$, then

$$\delta_i^0(\underline{z}_i) = \begin{cases} 1 & \text{if } p_i(\underline{z}_i) > cp_0(\underline{z}_i) \\ \lambda_0 & \text{if } p_i(\underline{z}_i) = cp_0(\underline{z}_i) \\ 0 & \text{if } p_i(\underline{z}_i) < cp_0(\underline{z}_i) \end{cases}$$

maximizes $\inf_{\underline{\theta}\in\bar{\Omega}} S(\underline{\theta},\delta)$ subject to $\sup_{\underline{\theta}\in\Omega} R(\underline{\theta},\delta) \le \gamma'$.

Example:

Let $X_{i1},\ldots,X_{in_i}$ be a random sample from $N(\theta_i,\sigma^2)$, $1 \le i \le k$, where σ^2 is known. Then $(\bar{X}_1,\ldots,\bar{X}_k)$ is a sufficient statistic for $\underline{\theta}' = (\theta_1,\ldots,\theta_k)$, where $\bar{X}_i = \sum_{\ell=1}^{n_i} X_{i\ell}/n_i \sim N(\theta_i,\sigma^2/n_i)$. Consider the transformations $g_c(\bar{x}_1,\ldots,\bar{x}_k) = (\bar{x}_1+c,\ldots,\bar{x}_k+c)$, then $\underline{Z}_i' = \{\bar{X}_i-\bar{X}_j;\ j \ne i,\ 1 \le j \le k\}$ is a maximal invariant. The induced group $\bar{G} = \{\bar{g}_c | \bar{g}_c(\theta_1,\ldots,\theta_k) = (\theta_1+c,\ldots,\theta_k+c)\}$ has maximal invariant $\underline{\tau}_i' = \{\theta_i-\theta_j;\ j \ne i,\ 1 \le j \le k\}$ and the distribution of $\underline{Z}_i$ depends only on $\underline{\tau}_i$. For any i, the joint density of $\underline{Z}_i$ is given by

$$P_{\underline{\theta}}(\underline{z}_i) = P_{\underline{\tau}_i}(\underline{z}_i) = (2\pi\sigma^2)^{-(k-1)/2}|\Sigma_i|^{-\frac{1}{2}}\exp\{-(\underline{z}_i-\underline{\tau}_i)'\Sigma_i^{-1}(\underline{z}_i-\underline{\tau}_i)/2\sigma^2\}$$

where

$$\Sigma_i = \begin{pmatrix} \frac{1}{n_i}+\frac{1}{n_1} & & \frac{1}{n_i} \\ & \ddots & \\ \frac{1}{n_i} & & \frac{1}{n_i}+\frac{1}{n_k} \end{pmatrix}_{(k-1)\times(k-1)}$$

"$\leftrightarrow$" ("$\updownarrow$") means that the ith row (ith column) is deleted.

$P_{\underline{T}_i}(\underline{z}_i)$ has SIP in $\underline{T}_i$, $1 \leq i \leq k$ and

$$P_i(\underline{z}_i)/P_0(\underline{z}_i) = \exp\{\Delta n_i(\sum_{j\neq i} z_{ij}n_j)/N\sigma^2 - \Delta^2(n_i - \frac{n_i^2}{N})/2\sigma^2\}$$

which is non-decreasing in $\underline{z}_i$ ($N = \sum_{i=1}^{k} n_i$). Furthermore,

$$P_i(\underline{z}_i)/P_0(\underline{z}_i) > c_i \quad \text{iff} \quad \bar{X}_i > \sum_{j\neq i} n_j\bar{X}_j / \sum_{j\neq i} n_j + d_i.$$

Thus

$$\delta_i^0 = \begin{cases} 1 & \text{if } \bar{X}_i \geq \sum_{j\neq i} n_j\bar{X}_j / \sum_{j\neq i} n_j + d_i \\ 0 & \text{if } \bar{X}_i < \sum_{j\neq i} n_j\bar{X}_j / \sum_{j\neq i} n_j + d_i, \end{cases} \tag{3.7}$$

if $R(\underline{\theta},\delta^0)$ is maximized at $\theta_1 = \theta_2 = \ldots = \theta_k$, then we have

(i) if $d_i = \Delta - \Phi^{-1}(\gamma)\sigma\sqrt{\frac{1}{n_i} + \frac{1}{\sum_{j\neq i} n_j}}$, then δ^0 defined by (3.7) minimizes $\sup_{\underline{\theta}\in\Omega} R(\underline{\theta},\delta)$ subject to $\inf_{\underline{\theta}\in\bar{\Omega}} S(\underline{\theta},\delta) \geq \gamma$.

(ii) If $\sum_{i=1}^{k} \Phi\left(\frac{-d_i}{\sigma\sqrt{\frac{1}{n_i} + \frac{1}{\sum_{j\neq i} n_j}}}\right) = \gamma'$ and $\Phi\left(\frac{\Delta - d_i}{\sigma\sqrt{\frac{1}{n_i} + \frac{1}{\sum_{j\neq i} n_j}}}\right) = \text{constant}$,

then δ^0 defined by (3.7) maximizes $\inf_{\underline{\theta}\in\bar{\Omega}} S(\underline{\theta},\delta)$ subject to $\sup_{\underline{\theta}\in\Omega} R(\underline{\theta},\delta) \leq \gamma'$.

In particular, if $n_1 = n_2 = \ldots = n_k = n$, then

$$\delta_i^0 = 1 \quad \text{iff} \quad \bar{X}_i \geq \frac{1}{k-1}\sum_{j\neq i} \bar{X}_j + d. \tag{3.8}$$

We know that $R(\underline{\theta},\delta^0)$ is maximized at $\theta_1 = \ldots = \theta_k$ iff $\inf_{\underline{\theta}\in\Omega} S(\underline{\theta},\delta^0) \geq \frac{k-1}{k}$ (see Bjørnstad (1981)). Therefore, if $\Phi(((k-1)/k)^{\frac{1}{2}}n^{\frac{1}{2}}d/\sigma) \leq \frac{1}{k}$ and

$d = \Delta - \sigma\Phi^{-1}(\gamma)n^{-\frac{1}{2}}(\frac{k-1}{k})^{-\frac{1}{2}}$, then δ^0 defined by (3.8) minimizes $\sup_{\underline{\theta}\in\Omega} R(\underline{\theta},\delta)$ subject to $\inf_{\underline{\theta}\in\bar{\Omega}} S(\underline{\theta},\delta) \geq \gamma$. If $\Phi(((k-1)/k)^{\frac{1}{2}}n^{\frac{1}{2}}d/\sigma) \leq \frac{1}{k}$ and $d = -\Phi^{-1}(\frac{\gamma'}{k})\sigma n^{-\frac{1}{2}}(\frac{k-1}{k})^{-\frac{1}{2}}$, then δ^0 defined by (3.8) maximizes $\inf_{\underline{\theta}\in\bar{\Omega}} S(\underline{\theta},\delta)$ subject to $\sup_{\underline{\theta}\in\Omega} R(\underline{\theta},\delta) \leq \gamma'$.

4. Optimal Selection Rules in Relation to Multiple Tests

Let $\underline{X} = (X_1,\ldots,X_k)$ be a random vector with probability distribution depending on a parameter vector $\underline{\theta} = (\theta_1,\ldots,\theta_k) \in \Omega$. Consider a family of hypothesis testing problems

(4.1) H_i: $\underline{\theta} \in \Omega_{0i}$ against K_i: $\underline{\theta} \in \Omega_i$

where $\Omega_{0i} = \Omega - \Omega_i$, $1 \leq i \leq k$, and Ω_i, $1 \leq i \leq k$, are defined as in Section 2. We know that $\Omega_0 = \bigcap_{i=1}^{k} \Omega_{0i}$. A test of the hypotheses (4.1) will be defined to be a vector $(\delta_1(\underline{x}),\ldots,\delta_k(\underline{x}))$, where the elements of the vector are ordinary test functions. When $\underline{x}$ is observed, we reject H_i with probability $\delta_i(\underline{x})$, $1 \leq i \leq k$. The power function of a test $(\delta_1,\ldots,\delta_k)$ is defined to be the vector $(\beta_1(\underline{\theta}),\ldots,\beta_k(\underline{\theta}))$, where $\beta_i(\underline{\theta}) = E_{\underline{\theta}}\delta_i$, $1 \leq i \leq k$. For $\underline{\theta} \in \Omega_i$, we know that $\beta_i(\underline{\theta})$ is the probability of a correct selection and $\delta_i(\underline{x})$ is the probability of selecting the best population π_i.

Let S_{γ_1} be the class of all tests $(\delta_1,\ldots,\delta_k)$ such that

$$(4.2) \qquad \sup_{\underline{\theta}\in\Omega_0} \sum_{i=1}^{k} E_{\underline{\theta}}\delta_i \leq \gamma_1.$$

Hence the expected subset size for the selection rule δ over Ω_0 in S_{γ_1} is less than or equal to γ_1. For each i, $1 \leq i \leq k$, we would, subject to (4.2), like to have $\beta_i(\underline{\theta})$ large when $\underline{\theta} \in \Omega_i$. For $\underline{\theta} \in \Omega_i$, if we make $\beta_i(\underline{\theta})$ large, then $\beta_j(\underline{\theta})$ will often have to be small for $j \neq i$, if (4.2) is to be satisfied.

Therefore, we will restrict attention to tests which

(1) maximize the minimum average power over Ω_i, $1 \le i \le k$, i.e. maximize

$\sum_{i=1}^{k} \inf_{\underline{\theta} \in \Omega_i} E_{\underline{\theta}} \delta_i(\underline{X})$ among tests δ in S_{γ_1}, or

(2) maximize the minimum power over Ω_i, $1 \le i \le k$, i.e. maximize

$\min_{1 \le i \le k} \inf_{\underline{\theta} \in \Omega_i} E_{\underline{\theta}} \delta_i(\underline{X})$ among tests δ in S_{γ_1} (see Spjøtvoll (1972) and Gupta and Huang (1977)).

As discussed in Section 2, we will assume that, for any i, the statistic $\underline{Z}_i = \{Z_{ij};\ j \neq i,\ 1 \le j \le k\}$ is sufficient invariant under a transformation group and has joint distribution which depends only on $\underline{\tau}_i$, say $P_{\underline{\tau}_i}(\underline{z}_i)$, with SIP in $\underline{\tau}_i$. Let $P_{\underline{\Delta}}(\underline{z}_i) = p_i(\underline{z}_i)$, $1 \le i \le k$, and $P_{\underline{\tau}_i}(\underline{z}_i) = P_0(\underline{z}_i)$ when $\tau_{ij} = \tau_{ii} =$ constant, $j \neq i$, $1 \le j \le k$.

Gupta and Huang (1977) have considered the first problem. In this section, we will consider the second problem.

Theorem 4.1. Suppose that for any i, $p_i(\underline{z}_i)/p_0(\underline{z}_i)$ is non-decreasing in $\underline{z}_i$. If δ^0 is given by

$$\delta_i^0(\underline{z}_i) = \begin{cases} 1 & \text{if } c_i p_i(\underline{z}_i) > p_0(\underline{z}_i) \\ \lambda_i & \text{if } c_i p_i(\underline{z}_i) = p_0(\underline{z}_i) \\ 0 & \text{if } c_i p_i(\underline{z}_i) < p_0(\underline{z}_i) \end{cases} \tag{4.3}$$

where c_i (> 0) and λ_i are determined by $\sum_{i=1}^{k} \int \delta_i^0 p_0 = \gamma_1$ and $\int \delta_i^0 p_i$ is independent of i. If $\sup_{\underline{\theta} \in \Omega_0} \sum_{i=1}^{k} E_{\underline{\theta}} \delta_i^0$ occurs at $\tau_{ij} = \tau_{ii} =$ constant, then δ^0 maximizes $\min_{1 \le i \le k} \inf_{\underline{\theta} \in \Omega_i} E_{\underline{\theta}} \delta_i$ among all rules δ in S_{γ_1}.

Proof. Since $p_i(\underline{z}_i)/p_0(\underline{z}_i)$ is non-decreasing in $\underline{z}_i$, then $\delta_i^0(\underline{z}_i)$ is non-decreasing in $\underline{z}_i$. By Lemma 3.1, for any $\underline{\theta} \in \Omega_i$,

$$E_{\underline{\theta}}\delta_i^0 \geq E_{\underline{\Delta}}\delta_i^0 = \int \delta_i^0 p_i$$

which is independent of i, by assumption.

Hence

$$\inf_{\underline{\theta}\in\Omega_i} E_{\underline{\theta}}\delta_i^0 = \int \delta_i^0 p_i. \tag{4.4}$$

Furthermore,

$$\begin{aligned}
&(\sum_{i=1}^{k} c_i)(\min_{1\leq i\leq k} E_{\underline{\Delta}}\delta_i^0 - \min_{1\leq i\leq k} E_{\underline{\Delta}}\delta_i) \\
&= \sum_{i=1}^{k} c_i(\min_{1\leq i\leq k} E_{\underline{\Delta}}\delta_i^0 - \min_{1\leq i\leq k} E_{\underline{\Delta}}\delta_i) \\
&\geq \sum_{i=1}^{k} c_i(\int \delta_i^0 p_i - \int \delta_i p_i) \\
&= \int \sum_{i=1}^{k} (\delta_i^0 - \delta_i)(c_i p_i - p_0) + \int \sum_{i=1}^{k} (\delta_i^0 - \delta_i) p_0 \\
&\geq 0, \text{ by definition of } \delta_i^0 \text{ and the fact} \\
&\sum_{i=1}^{k} \int \delta_i p_0 \leq \sup_{\underline{\theta}\in\Omega_0} \sum_{i=1}^{k} E_{\underline{\theta}}\delta_i \leq \gamma_1 = \sum_{i=1}^{k} \int \delta_i^0 p_0.
\end{aligned}$$

Hence

$$\min_{1\leq i\leq k} E_{\underline{\Delta}}\delta_i^0 \geq \min_{1\leq i\leq k} E_{\underline{\Delta}}\delta_i. \tag{4.5}$$

From (4.4) and (4.5), we have

$$\min_{1\leq i\leq k} \inf_{\underline{\theta}\in\Omega_i} E_{\underline{\theta}}\delta_i^0 = \min_{1\leq i\leq k} E_{\underline{\Delta}}\delta_i^0 \geq \min_{1\leq i\leq k} E_{\underline{\Delta}}\delta_i \geq \min_{1\leq i\leq k} \inf_{\underline{\theta}\in\Omega_i} E_{\underline{\theta}}\delta_i.$$

This completes the proof of the theorem.

Remarks:

(1) Theorem 4.1 implies Theorem 3.4. Since $\min_{1\le i\le k} \inf_{\underline{\theta}\in\Omega_i} E_{\underline{\theta}}\delta_i = \inf_{\underline{\theta}\in\bar{\Omega}} S(\underline{\theta},\delta)$

and if $\sup_{\underline{\theta}\in\Omega} R(\underline{\theta},\delta) \le \gamma_1$, then $\delta \in S_{\gamma_1}$.

(2) Theorem 4.1 can also be proved by using Lemma 3.2 and following the same arguments as in the proof of Theorem 3.4.

(3) If $c_i = c$, $1 \le i \le k$, then Theorem 4.1 follows from the theorem of Gupta and Huang (1977). Since

$$\min_{1\le i\le k} \inf_{\underline{\theta}\in\Omega_i} E_{\underline{\theta}}\delta_i^0 - \min_{1\le i\le k} \inf_{\underline{\theta}\in\Omega_i} E_{\underline{\theta}}\delta_i$$

$$= \frac{1}{k}\sum_{i=1}^{k} \inf_{\underline{\theta}\in\Omega_i} E_{\underline{\theta}}\delta_i^0 - \min_{1\le i\le k} \inf_{\underline{\theta}\in\Omega_i} E_{\underline{\theta}}\delta_i$$

$$\ge \frac{1}{k}\sum_{i=1}^{k} \inf_{\underline{\theta}\in\Omega_i} E_{\underline{\theta}}\delta_i^0 - \frac{1}{k}\sum_{i=1}^{k} \inf_{\underline{\theta}\in\Omega_i} E_{\underline{\theta}}\delta_i$$

$$\ge 0.$$

REFERENCES

1. Alam, K. (1973). On a multiple decision rule. Ann. Statist. 1, 750-755.

2. Bahadur, R. R. (1950). On a problem in the theory of k populations. Ann. Math. Statist. 21, 362-375.

3. Berger, R. L. and Gupta, S. S. (1980). Minimax subset selection rules with applications to unequal variance (unequal sample size) problems. Scand. J. of Statist. 7, 21-26.

4. Bjørnstad, J. F. (1981). A class of Schur-Procedures and minimax theory for subset selection. Ann. Statist. 9, 777-791.

5. Eaton, M. L. (1967). Some optimum properties of ranking procedures. Ann. Math. Statist. 38, 124-137.

6. Gupta, S. S. and Huang, D. Y. (1977). Some multiple decision problems in analysis of variance. Commun. Statist. A6(11), 1035-1054.

7. Gupta, S. S. and Huang, D. Y. (1980). A note on optimal subset selection procedures. Ann. Statist. 8, 1164-1167.

8. Gupta, S. S. and Huang, D. Y. (1981). Multiple statistical decision theory: recent developments. Lecture Notes in Statistics 6, Springer-Verlag, New York.

9. Lehmann, E. L. (1961). Some model I problems of selection. Ann. Math. Statist. 32, 990-1012.

10. Lehmann, E. L. (1966). On a theorem of Bahadur and Goodman. Ann. Math. Statist. 37, 1-6.

11. Nagel, K. (1970). On subset selection rules with certain optimizing properties. Ph.D. Thesis (Mimeo. Ser. No. 222), Dept. of Statist., Purdue University, West Lafayette, IN.

12. Spjøtvoll, E. (1972). On the optimality of some multiple comparison procedures. Ann. Math. Statist. 43, 398-411.

13. Studden, W. J. (1967). On selecting a set of k populations containing the best. Ann. Math. Statist. 38, 1072-1078.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER Technical Report #84-7	2. GOVT ACCESSION NO. AD-A140 180	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) ON THE OPTIMALITY OF SOME SUBSET SELECTION PROCEDURES		5. TYPE OF REPORT & PERIOD COVERED Technical 6. PERFORMING ORG. REPORT NUMBER Technical Report #84-7
7. AUTHOR(s) Lii-Yuh Leu		8. CONTRACT OR GRANT NUMBER(s) N00014-75-C-0455
9. PERFORMING ORGANIZATION NAME AND ADDRESS Purdue University Department of Statistics West Lafayette, IN 47907		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Washington, DC		12. REPORT DATE March 1984 13. NUMBER OF PAGES 14
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION. DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release, distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Subset selection, restricted minimax, multiple tests.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

As measures of goodness of a selection rule, usually two quantities, the probability of a correct selection and the expected size of the selected subset, are considered. Based on these two criteria, Gupta and Huang (1980) proved a theorem to derive a selection procedure with some optimality property. However, the theorem cannot be applied to the unequal sample sizes case. In this paper, we use a different method to generalize this theorem to the unequal sample sizes case. Also a dual problem is investigated. Also, we treat a selection procedure in terms of multiple tests. Based on this approach, we derive an optimality result.

DD FORM 1 JAN 73 1473

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

END

FILMED

5-84

DTIC